The pronunciation of voiced obstruents in L2 French: A preliminary study of Swiss German learners

Schmid, Stephan

Abstract: The present study examines how Swiss German learners cope with the contrast between voiced and unvoiced obstruents in L2 French. The feature [±voice]) is not exploited in Swiss German dialects, where pairs of obstruents sharing the same place and manner of articulation are basically differentiated in terms of longer or shorter duration (i.e., the feature [±tense]). Therefore, we expect that Swiss German learners of French would assimilate the non-native feature [±voice] to the native [±tense] contrast, due to the great similarity and the functional equivalence of the two features; devoicing is predicted to occur more often in universally preferred positions such as the prepausal context. The corpus consists of 20 sentences (containing 6 voiced obstruents in 6 different phonotactic contexts), which were read by 10 high school students. An acoustic analysis permitted to categorize the 340 tokens into three discrete types: fully voiced, fully unvoiced, partially voiced. Chi-square tests yielded significant effects of the factors "context", "segment" and "speaker" on the variable "voicing". In particular, speakers pronounced 58% of the intervocalic obstruents as fully voiced, whereas they devoiced 85% of the prepausal tokens (thus, revealing both L1-based and universally preferred patterns).

DOI: https://doi.org/10.1515/psicl-2012-0028

Originally published at:
Schmid, Stephan (2012). The pronunciation of voiced obstruents in L2 French: A preliminary study of Swiss German learners. Poznan Studies in Contemporary Linguistics, 48(4):627-659.
DOI: https://doi.org/10.1515/psicl-2012-0028

Poznań Studies in Contemporary Linguistics 48(4), 2012, pp. 627–659

doi:10.1515/psicl-2012-0028

THE PRONUNCIATION OF VOICED OBSTRUENTS IN L2 FRENCH: A PRELIMINARY STUDY OF SWISS GERMAN LEARNERS

STEPHAN SCHMID
University of Zurich
schmidst@pholab.uzh.ch

ABSTRACT

The present study examines how Swiss German learners cope with the contrast between voiced and unvoiced obstruents in L2 French. The feature [±voice]) is not exploited in Swiss German dialects, where pairs of obstruents sharing the same place and manner of articulation are basically differentiated in terms of longer or shorter duration (i.e., the feature [±tense]). Therefore, we expect that Swiss German learners of French would assimilate the non-native feature [±voice] to the native [±tense] contrast, due to the great similarity and the functional equivalence of the two features; devoicing is predicted to occur more often in universally preferred positions such as the prepausal context. The corpus consists of 20 sentences (containing 6 voiced obstruents in 6 different phonotactic contexts), which were read by 10 high school students. An acoustic analysis permitted to categorize the 340 tokens into three discrete types: fully voiced, fully unvoiced, partially voiced. Chi-square tests yielded significant effects of the factors "context", "segment" and "speaker" on the variable "voicing". In particular, speakers pronounced 58% of the intervocalic obstruents as fully voiced, whereas they devoiced 85% of the prepausal tokens (thus, revealing both L1-based and universally preferred patterns).

KEYWORDS: Obstruents; voicing; second language speech; French; Swiss German.

1. Introduction[1]

This study addresses the question of the extent to which Swiss German learners are able to produce voiced obstruents when they speak French. The paper is or-

[1] I would like to thank Claudia Mazza for her assistance in data collection, the participants of the study as well as Adrian Leemann for statistical advice. A first version of this work was presented at the Sixth International Symposium on the Acquisition of Second Language Speech "New Sounds

ganized as follows: the introductory section summarizes a number of studies on the pronunciation of L2 French (1.1) and describes the obstruents of the source and the target language, by illustrating the relevant subsets of the respective phoneme inventories, phonotactic constraints, and allophonic rules as well as the phonetic realization of stops and fricatives (1.2). Interpreting these descriptions in the light of current theories of second language speech (1.3), one may formulate two basic research hypotheses (which will be spelled out in more detail in 1.4): (i) Swiss German learners will, to a large extent, "assimilate" French voiced obstruents to the pattern of their native dialect; (ii) differences in the degree of voicing are expected to be found depending on the phonotactic context. Moving on to the empirical part of the study, Section 2 presents the informants who participated in the study (2.1), also documenting the procedures of data collection (2.2) and data analysis (2.3). Section 3 illustrates some of the data in a qualitative manner, i.e., by means of spectrograms (3.1), and then reports the results from a quantitative point of view, highlighting differences according to factors such as phonotactic context (3.2), place and manner of articulation (3.3), speaker (3.4), and speech rate (3.5); finally, a comparison is made with the realizations of voiced obstruents by a native speaker of French (3.6). The concluding remarks evaluate these results on the basis of the hypotheses stated above and discuss possible directions for future research (4.).

1.1. The pronunciation of French as a second language

Previous phonetic research has investigated several aspects of the pronunciation of French as a second language. Starting with prosody, Kaglik and Boula de Mareüil (2010) analyzed the French intonation of Polish learners and concluded that particular intonational groupings might derive not only from prosodic transfer from L1, but also from more general features of L2 prosody. Rhythm measures had been applied to the French of learners with six different native languages (i.e., Arabic, English, German, Italian, Portuguese, and Spanish) in a study by Vieru-Dimulescu and Boula de Mareüil (2006), who found, however, that vowel quality – in particular the pronunciation of /y/ – enabled the identification of the speakers' L1 much better than the rhythm measures; in a subsequent study with speakers of the same origin, vowel formants also revealed a

---

2010" in Poznań, 1–3 May 2010; cf. Schmid 2010). I am indebted to two anonymous reviewers who encouraged me to develop several aspects of this study further. Needless to say, I alone am responsible for any shortcomings or possible misinterpretations.

distinct influence of the relevant mother tongues (cf. Vieru-Dimulescu et al. 2007).

As regards the consonants of French, native speakers realize /p t k/ with a short-lag Voice Onset Time (VOT). Several researchers have investigated the production of unvoiced stops in L2 French by speakers of Germanic languages such as German, English and Dutch (cf., e.g., Künzel 1977; Flege 1987; Bongaerts 1999; Birdsong 2007), reporting a considerably long-lag VOT (i.e., a noticeable degree of aspiration); thus, the pronunciation of unvoiced stops seems to constitute a major source of difficulty at least for initial learners of French, who tend to transfer the VOT pattern of their L1 (German, English, and Dutch) to the second language.

From a review of the research literature one gains the impression that less attention has been paid to the topic of the present study, i.e., the production of voiced obstruents in L2 French. Still, Künzel's (1977) seminal study on German learners of French had also documented as principal sources of interference, besides the aspiration of voiceless stops, the devoicing of word-final stops and – quite interestingly – even the partial devoicing of word-internal stops (Künzel 1977: 176). Similar results have been found in the comparative study conducted by Vieru-Dimulescu, Boula de Mareüil, and Adda-Decker (2007: 2218): compared to French native speakers, German and in particular English learners realized a considerably lower percentage of voicing for the stops /b d g/ and a slightly lower voicing for the fricatives /v z ʒ/ (where the degree of voicing is defined in terms of duration, i.e., as the part of the segments for which F_0 values could be detected); conversely, Portuguese and Italian learners obtained degrees of voicing that were much closer to those found in speakers of the target language. Italian learners, however, may behave differently from native speakers with regard to a particular phonological rule of French, i.e., the assimilation of voiced/unvoiced consonant clusters (cf. 1.2), as has been demonstrated by Apolito and Gili-Fivela (2009).

Now, turning to the population under examination in this study, i.e., to Swiss German learners of French, we find some interesting observations in a contrastive grammar for language teachers. Hilty and Wüest (1985: 28) state that in Swiss German dialects the obstruents /b d g v z/ are produced without any participation of the vocal cords (cf. 1.2) and that therefore the pronunciation of voiced stops and fricatives requires special attention on the part of Swiss German learners of French; moreover, the authors maintain that in the case of /ʒ/ a "bad pronunciation" is particularly persistent. An experimental study with four speakers from Zurich (Horner 1989: 48) confirms – at least partially – the difficulty of this particular consonant: a waveform of the sentence *J'ai beau*

changer de logement (/ʒe bo ʃɑ̃ʒe də lɔʒmɑ̃/, 'I may well move to another house') does indeed show the devoicing of the postalveolar fricative in utterance-initial position ([ʃ̊]), but it also reveals that the same speaker is able to produce a fully voiced fricative in intervocalic position (*changer* [ʃɑ̃ʒe]). Thus, for the purpose of our study we will have to design a corpus that takes into account the different phonotactic contexts (cf. 2.2).

The importance of the phonotactic context has been demonstrated in an investigation of how Swiss German learners cope with voiced fricatives in another second language, namely English (Leemann 2011). Sixteen High School students from Berne read sentences containing fricatives in three different phonotactic positions, namely utterance-initially (##_), intervocalically (V_V) and utterance-finally (_##). The findings reveal differences depending on both place of articulation and context: on average, intervocalic realizations of /z/ and /ð/ are voiced, respectively, during 27% and 59% of the whole segment duration, whereas in the utterance-initial context the average voicing duration of the same consonants amounted to only 5% and 19% of the whole segment.

In order to understand the difficulties Swiss German learners may encounter in the pronunciation of French voiced obstruents better, we will now move on to a brief sketch of contrastive phonology. As we will see in the next subsection, the obstruent system of Swiss German dialects differs noticeably from the other languages considered so far, including Standard German.

1.2. Obstruents in French and in Swiss German dialects

If we look first at the phoneme inventory of the target language, we observe that the obstruent subsystem of French is rather simple and structurally coherent. Table 1, based on the illustration of the International Phonetic Alphabet provided by Fougeron and Smith (1999: 79), shows that there are only twelve obstruent phonemes, i.e., three pairs of stops – at the labial, coronal, and dorsal places of articulation – as well as three pairs of fricatives which are labiodental, dental, and postalveolar. In the core lexicon, French lacks affricates.

Table 1. Obstruent phonemes in French.

	Bilabial		Labiodental		Dental		Post-alveolar		Velar	
Plosive	p	b			t	d			k	g
Fricative			f	v	s	z	ʃ	ʒ		

The functional load of the phonemic voicing contrast is rather high, given that – in principle – it proves to be fully distinctive in word-initial, medial, and also in word-final position, as becomes clear from minimal pairs like /ʃu/ 'cabbage' ~ /ʒu/ 'to play (3rd pers. sg.)', /buʃe/ 'butcher' ~ /buʒe/ 'to move', /buʃ/ 'mouth' ~ /buʒ/ 'to move (3rd pers. sg.)'. Thus, a rule of word-final obstruent devoicing – present in several Germanic languages – is not part of French phonology; instead, a phonotactic constraint bans /z/ from the word-initial position. From a phonetic point of view, it is important to note that, in principle, "French voiced stops are typically voiced throughout" (Fougeron and Smith 1999: 80); the same – one could add – holds for fricatives as well. Empirical evidence for this claim comes, for example, from the acoustic measurements carried out in the above-mentioned (cf. 1.1) study by Vieru-Dimulescu et al. (2007).

Still, in French the feature [±voice] undergoes regressive assimilation if an obstruent is followed by another obstruent (cf. Léon 2007: 100). This phenomenon occurs both within words (e.g. [ɔpsɛʀve] 'to observe', [af̬gɑ̃] 'Afghan') and across word boundaries (e.g., /ʒə kʀwa/ [ʒ̊kɡ̊wa] 'I believe', /dɑ̃ sə baʀ/ [dɑ̃ s̬ baʀ] 'in this bar'); the latter examples show that such postlexical devoicing or voicing is likely to occur after schwa deletion. From a number of experimental studies, we know that these allophonic rules are – to some extent – variable, rather than categorical (Rigault 1970; Snoeren and Segui 2003; Darcy and Kügler 2007; D'Apolito and Gili Fivela 2009).

Contrary to French, Swiss German dialects lack voiced obstruents altogether, instead showing a binary opposition between so-called *fortis* and *lenis* consonants. Phonologically, the *fortis* vs. *lenis* contrast can be expressed by means of the feature [±tense] according to the proposal by Jakobson and Halle (1964: 100), who explicitly refer to the Swiss German system; recently, an alternative account of this type of contrast has been proposed for the Thurgovian dialect by Kraehenmann (2003), who opposes singleton to geminate consonants. In fact, it has been shown that the phonetic correlate of the *fortis* vs. *lenis* dichotomy basically rests on the amount of closure duration (Willi 1996; Nocchi and Schmid 2006). The following description follows the guidelines provided in the IPA illustration of the Zurich dialect (Fleischer and Schmid 2006), where *lenis* obstruents are transcribed with the diacritic for voicelessness (e.g., [b̥], [d̥], [ɡ̊]; [v̥], [z̥], [ɣ̊]).[2]

[2] It may be useful to specify that "Swiss German" is not a single language variety, but a commonly used cover term for the several dialects of German-speaking Switzerland. From a dialectological point of view, "alemannic" would be more appropriate (cf. Schubiger 1983); in fact, similar obstruent patterns occur in the neighboring dialects spoken in Germany and Austria.

Note that, in the present paper, the terms *fortis* and *lenis* are employed in accordance with traditional Swiss dialectology and in the narrow meaning with which they were introduced into phonetic theory by Winteler (1876: 21) and Sievers (1876: 65). A more recent terminological practice departs to some extent from the original definition, using *fortis* and *lenis* as cover terms for different consonantal contrasts that are related in one way or another to phonetic strength. In this latter view, both aspiration and voicing are seen as phonetic correlates of the *fortis* vs. *lenis* dichotomy; therefore, [b̥] and [b] would be labeled, respectively, as a "voiceless *lenis*" and a "voiced *lenis*" stop (cf., e.g., Künzel 1977: 8). In what follows, however, we will stick to the original definition (which is compatible with the Jakobsonian feature [±tense]) and simply refer to, e.g., [b̥] as a "*lenis* stop" and to [b] as a "voiced stop".

For the sake of clarity, it must be pointed out that in principle Swiss German obstruents are not voiced in any phonotactic or prosodic position, crucially differing in their phonetic substance from the obstruents of Standard German. While it is true that Standard German /b d g/ are often pronounced as voiceless in word-initial position (cf. Künzel 1977: 8; Kohler 1995: 158), they are normally voiced between two vowels; in addition, Standard German has voiced fricatives (Kohler 1995: 160), a consonant type which does not belong to the phonetics of Swiss German dialects.

Table 2. Obstruent phonemes in Swiss German dialects.

	Bilabial		Labiodental		Alveolar		Post-alveolar		Velar		Glottal
Plosive	p	b̥			t	d̥			k	g̊	
Fricative			f	v̥	s	z̥	ʃ	ʒ̊	x	ɣ̊	h
Affricate			pf		ts		tʃ		kx		

Table 2 shows that the phonological system of Swiss German dialects regularly exploits the feature [±tense] for plosive and fricative phonemes, whereas there is only a single series of affricates. Plosives display a broader phonotactic distribution than fricatives, which word-initially occur only as *lenes*, whereas tenseness is phonemically exploited in both word-internal and word-final contexts; for plosives, instead, the contrast is relevant in word-initial, word-internal, and word-final position. Accordingly, in Swiss German dialects there is no equivalent to the final devoicing process of Standard German (and in many oth-

er languages); indeed, both *fortis* and *lenis* obstruents can occur word-finally (Fleischer and Schmid 2006: 246).

Nevertheless, the *fortis–lenis* contrast is neutralized in another context, e.g., when a *lenis* plosive stands before a *fortis* plosive: in this case, a fortition of the *lenis* plosive occurs, e.g., /d̥ə z̥æb̥ tɒːg̊/ → [d̥ə z̥æp tɒːg̊] 'that day'. In addition, if a *lenis* plosive is followed by another *lenis* plosive, the result is again a *fortis* cluster, e.g., /heb̥ d̥i/ → [hep ti] 'hold tight!' (Fleischer and Schmid 2006: 248). Thus, two adjacent obstruents are treated differently in Swiss German than in French: in French, the first obstruent is regressively assimilated towards the voicing specification of the second obstruent (with two possible outcomes of the same rule), whereas in Swiss German a postlexical neutralization rule only yields one outcome, i.e., the unmarked *fortis* realization.

1.3. A quick look at some models of non-native speech perception and second language speech acquisition

As will become clear in the following sections (in particular 2.1), the character of the present study is observational rather than theoretical. It nevertheless fits some methodological standards of second language acquisition research (SLA), in that it applies an analytical approach, a deductive objective, and a relatively high degree of control and explicitness in data collection (cf. Seliger and Shohamy 1989: 22–41); within the qualitative–quantitative continuum of research methodologies proposed by Larsen-Freeman and Long (1991: 15), the design of our research might be collocated between the "focused description" and the "quasi-experimental" level. As regards second language theories, maybe an important issue should be clarified at this point: the fact that the preceding section (1.2) has been devoted to a sketch of contrastive analysis (limited to the phonological and phonetic description of obstruents in French and Swiss German dialects) by no means entails that we subscribe to the so called "Contrastive Analysis Hypothesis" (Lado 1957) which was proved decades ago to lead to wrong expectations about learner's errors. Rather, contrastive analysis is employed here as a research tool or a preliminary step in the preparation of the experimental design (cf. James 1980), but it does not suffice itself to generate predictions about learners' mental processes.[3]

[3] It seems that, in addition, more recent views of second language speech perception do implicitly involve some sort of contrastive analysis in their assumptions: for instance, basic linguistic terminology is employed in the claim that second language "learners' perception of L2 contrasts varies

Obviously, it lies outside the scope of this study to provide of an exhaustive account of current approaches in the fields of non-native speech perception and second language phonology – such as the "Speech Learning Model" (e.g., Flege 1995), the "Perceptual Assimilation Model" (e.g., Best 1995: 193–199; Best and Tyler 2007), or the "Ontogeny Phylogeny Model" (Major 2001: 80–134), to mention but three among the most influential approaches for the study of L2 speech. Nevertheless, we will briefly refer to some aspects of these models, which appear to be relevant for the topic of our empirical study.

Regarding the acquisition of voiced obstruents in L2 French, we have seen that in all L1 varieties considered so far (Dutch, English, Standard German, and the Swiss German dialects), pairs of homorganic obstruents are contrasted by a binary phonemic opposition, but all these languages differ from each other in the phonetic implementation of these contrasts. Such phonetic differences between the first and the second language might be less accessible to the phonological awareness of the learners, since experimental research on speech perception in a first language showed long ago that listeners have more sensitivity to differences between categories (phonemes) than to differences between categories (Liberman et al. 1957).

Turning now to second language speech, recent models have paid much attention to phonetic detail and to "perceived similarity/dissimilarity" between L1 and L2 speech sounds. Opposed to the "Contrastive Analysis Hypothesis," which stated that "those elements that are similar to the [learner's] native language will be simple for him" (Lado 1957: 2), the "Speech Learning Model" (SLM) claims that speech sounds which are only slightly different (i.e., "similar") are difficult to perceive and therefore difficult to acquire. More precisely, hypothesis 3 of the SLM states that "the greater the perceived phonetic dissimilarity between an L2 sound and the closest L1 sound, the more likely it is that phonetic differences between the sounds will be discerned" (Flege 1995: 239). Assuming that French voiced obstruents and Swiss German *lenes* are "similar" (in that the share most of the involved "phonetic categories"), the SLM would probably predict that a Swiss German learner of French will not discern the phonetic differences between two sounds such as [b] and [b̥]: therefore s/he will tend to perceive and pronounce L2 voiced obstruents as (voiceless) *lenes*. In other words, a Swiss German learner of French might classify voiced and *lenes* obstruents as "equivalent": "equivalence classification prevents adult L2 learners from establishing a phonetic category for similar but not new L2 phones" (Flege 1987: 50).

---

systematically according to L1 phonotactic, allophonic, and coarticulatory patterning" (Best and Tyler 2007: 19).

The "Perceptual Assimilation Model" (PAM) elaborated by Catherine Best (1995) does not constitute a theory of second language acquisition, but provides a number of predictions for cross-linguistic, non-native speech perception; it primarily focuses on the naïve, unexperienced listener who has at her/his disposition virtually no knowledge of the second language s/he is exposed to. Nevertheless, some implications of PAM for L2 research are discussed in Best and Tyler (2007: 24–32). One basic claim of PAM states that "naïve listeners, due to their native language experience, are likely to perceptually *assimilate* the non-native phone to the most articulatory-similar native phoneme" (Best and Tyler 2007: 22). For hypothetical naïve Swiss German listeners of French, PAM would probably predict that these listeners assimilate French voiced plosives and fricatives (i.e., the non-native phones) to the homorganic Swiss German *lenes* plosives and fricatives. In other words, for an unexperienced Swiss German ear, a French voiced plosive or fricative would constitute "an acceptable but not ideal exemplar" of the corresponding "native segmental category" (Best 1995: 194), i.e., the homorganic *lenis* plosive or fricative. Thus, for a non-native listener, PAM also seems to postulate an "equivalence classification", quite similarly to SLM. Nevertheless, for second language listeners PAM goes one step further, maintaining that "it is a the phonological level that listeners may identify L1 and L2 sounds as functionally equivalent" (Best and Tyler 2007: 25). Considering that there are many similar minimal pairs in French and in Swiss German that do precisely exploit the features [±voice] and [±tense] (cf., e.g., French /paʀ/ 'part' ~ /baʀ/ 'bar' vs. Zurich German /pɒːr/ 'pair' vs. /b̥ɒːr/ 'bar'), such functional equivalence may strongly bias the Swiss German listeners' perception of French voiced obstruents.

The role PAM ascribes to phonology leads us to stress the fact that SLM and PAM are characterized by a number of differences, but they also share some commonalities (cf. Best and Tyler 2007: 24–26). Differences focus on theoretical assumptions about human perception, the primitives of speech perception, the role of phonology as opposed to phonetics, and the particular type of listener the models focus on. PAM is theoretically grounded in the philosophy of "direct realism" which claims that "perceivers gain direct information from the world" (Best 1995: 173), whereas SLM seems to posit the existence of "long-term memory representations called phonetic categories" (Best and Tyler 2007: 24), which seem to constitute the primitives of speech perception; instead, PAM follows the theory of Articulatory Phonology in assuming "gestures" as phonological primitives (Best 1995: 187–193). Phonological structures are less important for SLM, whereas they are taken into account by PAM (cf. above). Finally, as has already been mentioned, PAM is basically a model of perception, whereas

the ultimate goal of SLM is to explain second speech learning. Among the commonalities we might recall that both models attribute a crucial role to perception and in particular to the perceived similarity/dissimilarity of L1 and L2 speech sounds; in a sense, both models are also strongly L1-biased.

A somewhat different stand is taken by the "Ontogeny Phylogeny Model" (Major 2001), which more explicitly takes into account fundamental notions of standard second language acquisition theory such as "interlanguage", "input", and the like;[4] moreover, the model addresses not only the second language speech of individuals (ontogeny), but it also considers more general linguistic phenomena such as language contact and change (Major 2001: 81). In a nutshell, the "Ontogeny Phylogeny Model" (OPM) states that during the acquisition of an L2 sound system "transfer processes decrease over time, while developmental processes increase and then decrease" (Major 2001: 80). Besides transfer from L1, the model also focuses on the developmental path towards the target language as well as on the emergence of universal patterns of sound structure (Major 2001: 6), among which we should mention the devoicing of word-final obstruents (cf. Major 2001: 4): indeed, this natural process occurs not only as a result of L1 transfer (as in the case of German learners of French mentioned above; cf. 1.1), but it also shows up through the adaptation of loanwords in languages which lack final obstruents, as was already argued by the founders of the theory of "Natural Phonology" (Donegan and Stampe 1979: 132–133).

The observation that German learners of English tend to devoice word-final obstruents is one piece of evidence also adduced in favor of the "Markedness Differential Hypothesis" (MDH), originally formulated by Eckman (1977).[5] This hypothesis draws on language typology (and in particular on implicational universals) in order to predict difficulties in second language acquisition: typologically marked structures, e.g., the voice contrast in syllable coda position, are

[4] To be precise, the proponents of PAM also recall the term "interlanguage", but as becomes clear from the following quote, they use it in a rather idiosyncratic and somehow reductive manner: "Numerous L2 speech researchers have posited that a learner's L1 and L2 phonological systems are not completely separate, but are instead situated within an encompassing interlanguage" (Best and Tyler 2007: 18). Compare this interpretation with a "state of the art" definition of the notion of "interlanguage" (Gass and Selinker 2008: 14): "The basic assumption in SLA research is that learners create a language system, known as an interlanguage (IL). This system is composed of numerous elements, not the least of which are elements of the NL (native language) and TL (target language). There are also elements of the IL that do not have their origins in either the NL or the TL. These latter are called new forms and are the empirical essence of interlanguage. What is important is that the learners themselves impose structure on the available linguistic data and formulate an internalized system (IL)".

[5] For a similar hypothesis stated in terms of linguistic naturalness, cf. Schmid (1997).

assumed to be more difficult to acquire (if they are absent in the L1) than typologically unmarked structures, e.g., the voice contrast in syllable onset position (Eckman 2008: 97).[6] It is not clear what kind of prediction this hypothesis would yield for the case of Swiss German learners of French. It appears that languages which contrast homorganic obstruents on the basis of the feature [±tense] are relatively rare, if compared with the many phonological systems that exploit the feature [±voice]. Therefore, on the basis of considerations of typological markedness (in the sense of cross-linguistic frequency), it should be easy for Swiss German learners to acquire the voice contrast in French. Nevertheless, they might also not perceive the difference between "voiced *lenes*" and "voiceless *lenes*" – because of the phonetic similarity of these two categories; this would mean that the attraction of perceptual assimilation (as claimed by models like SLM and PAM) overrides the predictive power of the MDH. Nevertheless, the MDH makes a strong point with regard to phonotactic contexts: voicing is more likely to occur in intervocalic position than word-finally or utterance-finally.

Given the growing body of knowledge about phonological typology (Maddieson 1984, 2011), one might even go one step further and investigate how the cross-linguistic frequency of the voice contrast is related to manner and place of articulation, and what kind of hypotheses for second language acquisition can be formulated on this ground. For instance, it appears that the voice contrast is more often exploited in stop systems than in fricative systems (Maddieson 1984: 35, 45);[7] moreover, within the series of voiced stops, /g/ is the phoneme that is most often absent in phoneme inventories, due to articulatory – or more precisely: aerodynamic – reasons (Ohala 1992).

Let us now briefly bring together the issues relevant to our study which have emerged so far from three types of source, i.e., from (i) empirical research on the pronunciation of French as a second language, (ii) a contrastive analysis of the French and Swiss German obstruent systems, and (iii) models of non-native speech perception and second language speech acquisition.

---

[6] Again, it lies outside the scope of this observational study to discuss in detail the many views of "markedness" that have been proposed in the history of linguistic theory from the Prague School to Optimality Theory. Instead, we refer to a long standing tradition in SLA research which has fruitfully investigated parallels between developmental processes in interlanguages and insights gained from language typology and universals (e.g. Rutherford 1984). Typological markedness – in particular cross-linguistic frequency of forms and structures as well as implicational universals – has in fact been acknowledged as a "research tool in the study of second language acquisition" (Hyltenstam 1990).

[7] In the UPSID database, the ratio between voiced and unvoiced segments amounts to 0.63 in the case of plosives and to 0.43 in the case of fricatives.

1.4. Research questions and hypotheses

The main purpose of this study is to investigate how speakers of Swiss German cope with the voice contrast in French obstruents. The anecdotal evidence provided by Hilty and Wüest (1985) and Horner (1989) suggests that these learners will tend to replace voiced obstruents with *lenis* obstruents (or, to use another terminology, they will replace "voiced *lenis*" obstruents with "unvoiced *lenis*" obstruents). This general hypothesis is further motivated by both the contrastive description of the language pair under investigation (cf. 1.2) and by general reflections about second language speech perception, such as those put forth by the "Speech Learning Model" and the "Perceptual Assimilation Model" (cf. 1.3). The structural equivalence of the two features [±voice] and [±tense] as well as their phonetic similarity would indeed foster this hypothesis.

Nevertheless, we can expect to find some variability in the realization of voiced French obstruents according to at least three factors. A first source of variability might be related to intersubject differences (due to personal factors such as the language biography or the learner's aptitude in pronouncing a new language). A second source of variability which needs to be controlled for derives from the phonotactic and prosodic context in which the sounds under analysis are uttered; both previous studies on French as a second language (cf. 1.2) and more theoretical considerations about typological markedness and/or universal naturalness suggest that voicing will be mastered to some extent in a word-internal, intervocalic context. Considering the native system of Swiss German learners, final obstruent devoicing need not necessarily apply at the word-level, but it can be expected to occur in prepausal position. Finally, it is worthy of investigation what kind of variability can be found with regard to the manner and place of the obstruents elicited from the learners.

In sum, at the beginning of this study we formulated two main hypotheses: (i) Swiss German learners of French will tend to pronounce "voiced *lenis*" obstruents as "voiceless *lenis*" obstruents; (ii) devoicing is expected to be more frequent in preconsonantal and utterance-final contexts than in intervocalic position.[8] If we now compare these hypotheses with the models discussed in 1.3, we do find certain coincidences. In our understanding, hypothesis (i) would in fact be supported both by the "Speech Learning Model" (SLM) and the "Perceptual

[8] Note that quite similar hypotheses had already been formulated by Schubiger (1983: 25): "Thus Swiss speakers of English – and of French – are apt to use the weak, voiceless consonant for the weak, voiced one. The sounds in question are the plosives b, d, g, and the fricatives z, ʒ, dʒ (*sic*), v at the end of words, and by analogy ð".

Assimilation Model" (PAM), mainly on the grounds of perceived similarity and equivalence classification (i.e., functional equivalence in the case of PAM). Probably PAM would also support the second hypothesis, since it emphasizes "allophonic or phonotactic or sub-phonemic (phonetic) influences of the L1 on perception of L2 and nonnative speech" (Best and Tyler 2007: 22). Note that the second hypothesis would also appear to be supported by the "Markedness Differential Hypothesis", given the cross-linguistic frequency of voiced obstruents in intervocalic position and the cross-linguistically common phenomenon of word- and utterance-final obstruent devoicing. Now, moving beyond the two basic hypotheses stated in the introduction, the consideration of the cross-linguistic frequency of certain consonant types leads us to test two additional hypotheses: (iii) fricatives might undergo devoicing more frequently than stops; (iv) among stops, /g/ will be the segment with the highest percentage of devoicing. Anticipating information that will be provided in the next section, the linguistic repertoire of some participants in our study will suggest a fifth hypothesis: (v) learners with a particular bilingual background (e.g., Swiss German and Italian) could be more prone to realize voiced obstruents than monolingual speakers of Swiss German (cf. 2.1 and 3.4).

2. Data and methods

In order to test the five afore-mentioned hypotheses, a corpus of read speech was collected among Swiss German learners of French. The same speech material was also read by a native speaker of French. Subsequently, the data was analyzed at the Phonetics Laboratory of the University of Zurich. The following subsections will present the speakers who participated in the experiment (2.1), the structure of the elicited corpus (2.2), and the data analysis procedure (2.3).

2.1. Speakers

The participants in the study were 10 students who attended a public high school in the city of Zurich; at the time of the recording (in the year 2007), the students were aged sixteen or seventeen and had experienced seven years of formal instruction in French (on average three hours per week). Their level of L2 competence can be described as "upper intermediate" or B2 in terms of the "Common European Framework of Reference of Languages" (CEFR 2001).

The fact that we are dealing here with (instructed) foreign language acquisition (FLA) rather than with (spontaneous) second language acquisition (SLA) raises the question as to what extent the participants can be considered as "experienced" or "inexperienced" learners; however, even if FLA is assumed to be "a fairly impoverished context for L2 learning", "FLA listeners, like L2 listeners but unlike monolinguals, have exposure to the target language" (Best and Tyler 2007: 19).

Both German and French are national languages in Switzerland, together with Italian and Romansh; in the 2000 census, about two third of the population declared German as the dominant language, whereas about one fifth of the inhabitants indicated French as their main language (Manno 2003, 2007). In most cantons of German-speaking Switzerland, French is an obligatory subject in public schools, and for many decades it was the first L2 children were taught. In 1997, however, the canton of Zurich anticipated the teaching of English as a foreign language in the curriculum of primary schools; thus, the participants in the present study had started to learn English before French (Manno 2007). Moreover, many high school teachers report that the English language is more popular among the youngsters, who are less motivated to learn French than English.

Despite a well-established myth, Switzerland is not a multilingual country in the sense that all its inhabitants speak several of the four national languages fluently. Nevertheless, individual bilingualism does occur in certain families as a consequence of immigration. So, although all the participants in our study speak Zurich German as their main language in everyday communication, some of them are bilingual speakers of another language. Among the 10 students who participated in the task, 6 had only Swiss parents (Lis, Raf, Ang, Van, Mir, Ser). The remaining 4 students had varied language backgrounds: both parents Italian (Dav), one parent Italian (Seb, Mel), one parent Finnish (Lar). The presence of Italian in the repertoire of some participants might intervene in their pronunciation of French (cf. hypothesis (v) stated in 1.4), given that Italian learners had been reported to produce a relatively high degree of voicing in L2 French (Vieru-Dimulescu et al. 2007: 2218).

2.2. Speech materials

In order to test the hypotheses spelled out in 1.4, a corpus of read speech was collected. The speech materials consist of 20 simple declarative and interrogative sentences, the content of which refers to objects and events of everyday life

(cf. Appendix); the French teacher confirmed that the students were familiar with all lexical items. Carrier sentences were avoided not only in order to obtain relatively "natural" read speech, but also because the target words had to occur in different contexts (two of them being determined by the position of the word within the utterance). In fact, the corpus contains the 6 voiced obstruents /b d g v z ʒ/ in 6 different phono(syn)tactic contexts: (i) ##_V, (ii) V#_V, (iii) V_V, (iv) V_C[+voice], (v) V_C[−voice], (vi) V_##. In the first context (##_V), the segment examined occurs utterance-initially, whereas in the second case (V#_V) the word-initial consonant is preceded by a word-final vowel. Within the word, the segment also occurs intervocalically (V_V) and before a voiced or an unvoiced consonant (i.e., V_C[+voice] and V_C[−voice]); the last context contains the consonant in a word-final and utterance-final (prepausal) position (V_##). All in all, the sentences contain 34 target words according to the different segment types under analysis: 6 consonants multiplied by 6 phonotactic contexts minus 2 contexts, given that /z/ is not allowed word-initially. Since different phonotactic and prosodic contexts were required, some sentences contain more than one target word (moreover, several words occur in more than one sentence); thus, 20 sentences were enough to elicit all the segments needed for the purpose of the study.

Table 3 illustrates the 34 target words. Numbers between brackets refer to the sentences listed in the Appendix (which also provides a phonemic transcription and an English translation of the target sentences). The asterisk * indicates that in a particular context a given consonant is phonotactically illegal.

Table 3. Target words according to segment type and context.

	b	d	g	v	z	ʒ
##_V	beaucoup (5)	dans (6)	gardez (17)	voulez (4)	*	je (1)
V#_V	boire (4)	dois (2)	garçon (9)	vous (4)	*	je (2)
V_V	auberge (3)	adore (8)	bagages (17)	souvent (1)	maison (6)	manger (2)
_#C[+voice]	robe (20)	aide (9)	vagues (8)	trouve (3)	cause (16)	auberge (3)
_#C[–voice]	robe (19)	aide (10)	vagues (13)	lève (18)	chemise (15)	mange (7)
_##	robe (14)	aide (11)	blague (12)	grève (16)	chose (2)	plage (1)

The recordings took place in the library of the high school building by means of a digital field recorder and a highly directional supercardioid microphone (fre-

quency response: 50 Hz–20 kHz ± 2.5dB); a sampling rate of 44.1 kHz and a quantization of 16 bit were employed. During the recording session, the participants had to read the sentences twice; between the two readings, they were engaged in a short informal conversation regarding topics such as spare time and holidays. The purpose of the conversation was twofold: on the one hand, it was meant to provide a break between the two readings, making the task less boring, while on the other hand it permitted us to elicit two speaking styles, namely reading vs. semi-spontaneous speech. A possible hypothesis to test would state that devoicing is more frequent in the conversation than in the reading data; however, for the present study the conversation data have not been analyzed, such that for the time being a comparison between the two speaking styles is not possible.

2.3. Data analysis

As regards the acoustic analysis of the 340 tokens (the 34 segment types mentioned above read by 10 participants), waveforms and spectrograms were inspected manually using the software *Praat* (Boersma and Weenink 2012). For each token, the duration of the periodic signal was measured taking into account three different acoustic cues, i.e., the presence of glottal pulses, the presence of a fundamental frequency contour, and the shape of the waveform (in most cases, there was a correspondence between these cues, among which, however, the presence of an F_0 contour appeared to be the least reliable).

Considering that the spectrographic inspection yielded three different types of fricatives (i.e., fully voiced, fully voiceless, and partially voiced; cf. 3.1), the amount of voicing time was noted for all segments under analysis. For instance, a fully voiced obstruent was assigned a score of 100%, while a fully devoiced obstruent obtained a score of 0%; if a consonant was partially voiced (at the beginning or at the end), the percentage of the duration of the voiced part over the duration of the entire segment was retained (e.g. 30%). On the whole, however, the realizations showed a tendency towards the voiced or the unvoiced pole of the continuum and yielded – at least for different contexts – a bimodal distribution; thus, a quantitative analysis in terms of three discrete categories (voiced, partially voiced, unvoiced) was considered to be sufficient. Given that the data had been categorized into a nominal scale, the effect of the three factors "context," "segment," and "speaker" on the variable voicing was calculated by means of a chi-square test.

3. Results

In the following sections, the recorded obstruents are analyzed from different point of views. First, a spectrographic illustration is provided of three types of fricatives – i.e., fully voiced, partially voiced, and fully devoiced (3.1). Second, the number of voiced realizations is illustrated according to the phonotactic contexts in which the segments were uttered (3.2) and according to the place and manner of the segments (3.3). Subsequently, differences in pronunciation accuracy among the 10 learners are shown (3.4), also with regard to speech rate (3.5); finally, a comparison with the reading of a native speaker is made (3.6).

3.1. Voiced, partially voiced and fully devoiced fricatives

Figure 1 shows the spectrogram of a fully voiced fricative in intervocalic position, as the speaker Dav pronounces the word *manger* /mɑ̃ʒe/ 'to eat' (sentence 2).

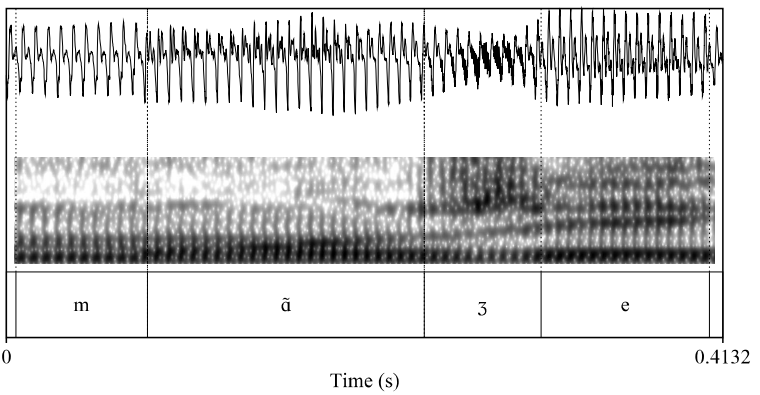

Figure 1. Waveform and spectrogram of the word *manger* /mɑ̃ʒe/ 'to eat', pronounced as [mɑ̃ʒe] by the speaker Dav.

As is evident from the periodic oscillation in the waveform as well as from the continuous voice bar in the lower part in the spectrogram, the vocal folds are vibrating throughout the articulation of the fricative [ʒ].

In the word-final context of sentence 1, the same speaker (Dav) also produces a realization of the fricative /ʒ/ which is fully devoiced (i.e., [ʃ]), as becomes clear from Figure 2; thus we may expect to find some degree of intrasubject variation according to the phonotactic context.

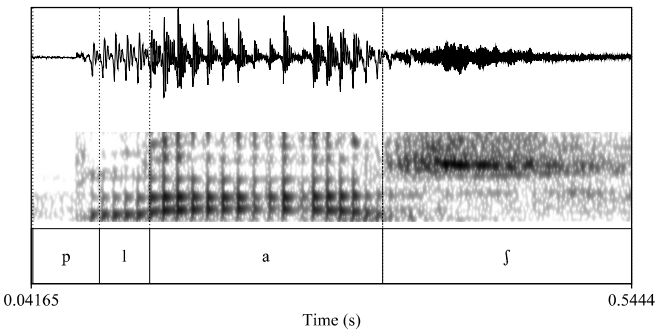

Figure 2. Waveform and spectrogram of the word *plage* /plaʒ/ 'beach', pronounced as [plaʃ] by the speaker Dav.

Still, a third type of realization – namely, a partially voiced fricative – appears in Figure 3, which illustrates the pronunciation of the word *manger* 'to eat' by the speaker (Lis): here, /ʒ/ loses its periodicity in the middle of the articulation, the duration of the voiced part amounting to 50% of the whole segment.

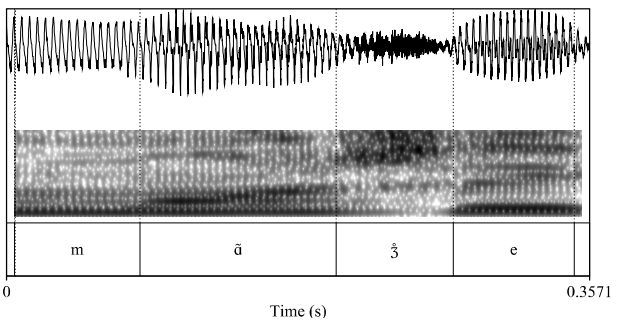

Figure 3. Waveform and spectrogram of the word *manger* /mɑ̃ʒe/ 'to eat', pronounced as [mɑ̃ʒ̊e] by the speaker Lis.

Out of the 340 tokens, 106 are fully voiced, 62 partially voiced, and 172 totally unvoiced. Within the intermediate category of the partially voiced obstruents, the periodic part ranges from 5% to 91% of the segment duration; the rather high dispersion is expressed by the standard deviation of 43 (mean: 48, median: 53). At this point, the question arises to what extent voicing is determined by phonotactic contexts and/or by individual differences among speakers.

3.2. Realizations of voiced obstruents according to phonotactic contexts

Figure 4 illustrates how the phonemically voiced obstruents of French are pronounced by our speakers. The diagram plots on the Y-axis the number of tokens belonging to the three categories of realizations (unvoiced, partially voiced, voiced), and on the X-axis the six different phonotactic contexts.

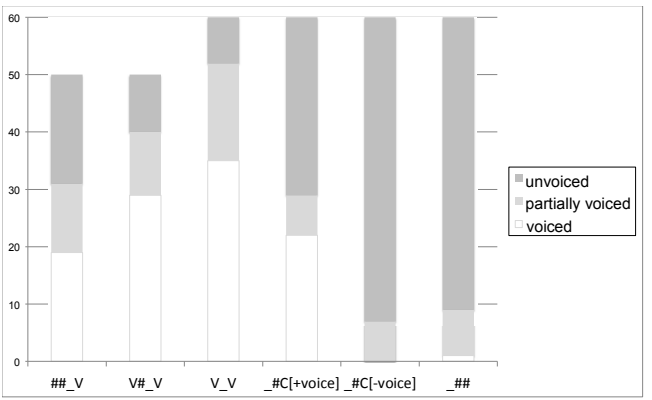

Figure 4. Realizations of French voiced obstruents according to phonotactic contexts.

The graph shows that the amount of voicing of the analyzed obstruents is indeed affected by their position in the sound chain. In particular, it appears that in word-internal intervocalic position (V_V) speakers attain the highest accuracy (35 voiced, 17 partially voiced, 8 unvoiced). The V#_V-column shows that the

presence of a word boundary has virtually no impact on the pronunciation of the obstruents (29 voiced, 11 partially voiced, 10 unvoiced); in both the V_V and the V#_V context, the percentage of target-like realizations amounts to 58%. This finding is far from surprising, as word boundaries do not play any substantial role in the phonology of either French or Swiss German (quite differently from Standard German). As expected, we find less voicing in the utterance-initial context (19 voiced, 12 partially voiced, 19 unvoiced), and the number of voiced tokens decreases further before a voiced consonant (_#C[+voice]), where we find only 22 voiced obstruents as opposed to 7 partially and 31 completely unvoiced obstruents; in this case, our speakers behave quite differently from the Italian learners analyzed by D'Apolito and Gili Fivela (2007), who were indeed able to pronounce sequences of two voiced obstruents. In a sense, the devoicing of a voiced obstruent before another voiced obstruent is in agreement with the phonotactics of Swiss German, where two adjacent *lenes* undergo fortition, but it is at odds with the native-like pronunciation of these consonant clusters. Instead, the phonologies of Swiss German and French go hand in hand when an underlying voiced (or *lenis*) obstruent is followed by an unvoiced (or *fortis*) obstruent; thus, the results of the #C[−voice] context come as no surprise: 0 voiced, 7 partially voiced, 53 unvoiced. Finally, the heaviest "violation" of French pronunciation rules is achieved in the last context: utterance-finally, only one obstruent is realized as voiced, against 8 partially voiced and 51 unvoiced obstruents (=85%). We interpret this finding as a reflection of a universal (articulatorily determined) preference towards devoicing in prepausal position.

As becomes clear from the descriptive statistics, the distribution of the three types of phonetic realization (voiced, partially voiced, unvoiced) is strongly determined by the phonotactic context of the segments under examination. Indeed, a chi-square test of independence confirmed that there are highly significant differences in the relative proportions of obstruent voicing between the six phonotactic contexts (X^2 (10, 340) = 129.0, $p < 0.0001$).

3.3. Realizations of voiced obstruents according to manner and place of articulation

A previous study on obstruent voicing in L2 French had shown that English and German learners performed slightly better with stops than with fricatives (Vieru-Dimulescu et al. 2007: 2218); this finding would be in line with predictions derived from cross-linguistic frequency, given that the languages of the world exploit the feature [±voice] to differentiate stops more often than frica-

tives (cf. 1.3). Hilty and Wüest (1985: 28) maintained that /ʒ/ would rarely be pronounced correctly by Swiss German speakers, whereas phonemic typology (grounded on considerations of articulatory ease) would point to /g/ as a marked – and therefore difficult – segment.

Now turning to our data, we find considerable differences in voicing between the 6 segments /b d g v z ʒ/. The prediction that fricatives would be more difficult to undergo voicing is not borne out by the production of our speakers. On the contrary, /v/ happens to be the segment that is most frequently realized as voiced (32 voiced, 8 partially voiced, 20 voiced), whereas the other two fricatives obtain lower scores, as can be seen from the columns of /z/ (7 voiced, 15 partially voiced, 18 unvoiced) and /ʒ/ (14 voiced, 14 partially voiced, 32 unvoiced). Note that the particularly "bad pronunciation" of /ʒ/ on the part of Swiss German learners of French (Hilty and Wüest 1985: 28) is not confirmed by our speakers. Thus, in the case of hypothesis (iii), cross-linguistic frequency does not suffice as a predictor for pronunciation difficulties in a second language; in fact, from a strictly articulatory or aerodynamic perspective it is not clear why fricatives should be more prone to devoicing than stops.

Instead, hypothesis (iv) – regarding the special status of the velar voiced stop, both from a typological and an aerodynamic point of view – appears to be corroborated by the histogram in Figure 5: indeed, /g/ has only 11 voiced realizations (8 partially voiced and 41 unvoiced), whereas the voiced tokens are more numerous both in the case of /d/ (23 voiced, 11 partially voiced, 26 unvoiced) and /b/ (19 voiced, 6 partially voiced, 35 unvoiced).

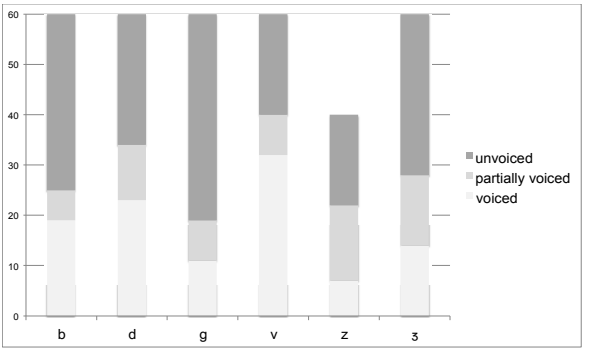

Figure 5. Realizations of French voiced obstruents according to manner and place of articulation.

Needless to say, these results are to be interpreted with caution, given the rather low number of tokens. Nevertheless, a chi-square test of independence indicates that the difference between the 6 obstruents in the relative proportions of voicing is significant (X^2 (10, 340) = 38.9, $p < 0.0001$).

3.4. Realizations of voiced obstruents according to speakers

Due to the phenomenological variability inherent to the interlanguage continuum (Tarone 1988), variation in a second language is normally larger than in first language data. The general problem of variation in second language acquisition is not at stake in this study, since the speech materials have been tightly controlled and the population is rather homogeneous in terms of age, type of input, and years of study of the foreign language. Still, it is reasonable to expect to find a certain amount of intersubject variability, due to individual factors such as language aptitude or linguistic biography (cf. Skehan 1989). In our case, it would be interesting to verify the extent to which the performance of the learners is influenced by their mastery and use of languages that do have fully voiced obstruents, such as Italian (cf. 2.1).

Figure 6 illustrates the individual differences observed among the participants of our study, as far as the realization of voiced obstruents is concerned. In the histogram, speakers are ordered from left to right according to the number of voiced obstruents, which gradually decreases from 16 (in Ang and Dav) towards 4 (Lar). As a general tendency, the number of unvoiced obstruents increases in indirect proportion to the number of voiced obstruents, with the exception of the speakers Mir and Lis, who realize a greater number of partially voiced occurrences (12 and 13 tokens).

Note that three speakers in the left part of the diagram (Dav, Seb, Mel) have a bilingual language background with Italian, a language which has fully voiced obstruents in all phonotactic contexts. It has been reported that voiced obstruents are indeed retained in the Italian spoken by immigrant children in Switzerland, who display only slight influences of the Swiss German dialect on the pronunciation of their heritage language (Schmid 2005). We can therefore interpret the relatively high number of voiced tokens realized by the three Swiss-German/Italian bilinguals as an influence of one of their two first languages on the production in L2 French. Now, it is also interesting to note that we find another bilingual at the righthand pole of the continuum displayed in Figure 6: the second native language of the speaker Lar is Finnish. From a description of Finnish phonology (Suomi et al. 2008: 25) we learn that the only voiced obstru-

ent phoneme appears to be /d/, even if "the plosives can be partly or fully voiced in fast and careless speech" (Suomi et al. 2008: 27). Considering that the speech rate of Lar is rather slow (cf. 3.5), we may again interpret the very high amount of unvoiced tokens in the sense that both native languages of this speaker act against the acquisition of voiced obstruents in a second language.[9]

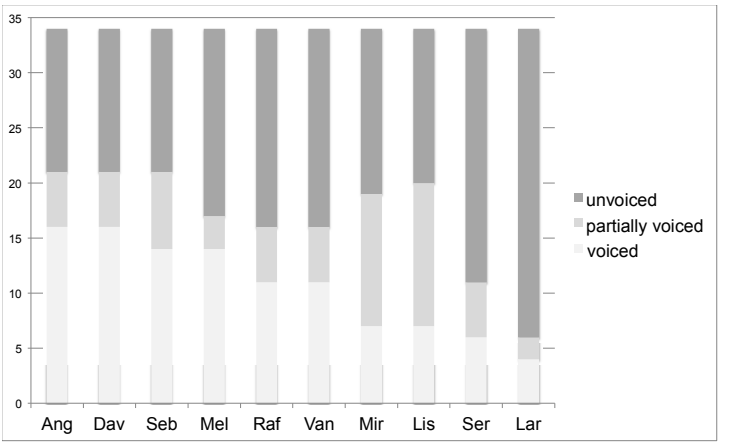

Figure 6. Realizations of French voiced obstruents according to speakers.

Individual differences in obstruent devoicing were predicted by the fifth hypothesis stated in 1.4 (at least as far as bilingual participants are concerned). From the more general point of view of statistical analysis, a chi-square test of independence reveals a highly significant effect of the factor "speaker" (X^2 (10, 340) = 47.7, p = 0.0002).

[9] One of the two anonymous reviewers proposed that we treat the Italian bilinguals separately in all the analyses and exclude the bilingual with Finnish altogether. It is true that the issue of bilingualism was originally not part of the research design; therefore, the sample of our participants exhibits a certain degree of heterogeneity which is at odds with the intended deductive make-up of the experiment. Given the preliminary and explorative nature of our study, we nevertheless refrained from adopting this recommendation, assuming that the results presented here may nevertheless be of some interest, in particular for future research on voiced obstruents in bilingual speakers.

3.5. Realizations of voiced obstruents and speech rate

One aspect by which individuals' speech may differ is, obviously, speech rate. Somehow surprisingly, though, previous research on voiced obstruents did not take this factor into account (cf. 1.1) and, accordingly, no specific hypothesis regarding speech rate was formulated in this study (cf. 1.4). Nevertheless, it is worthwhile having a closer look at speech rate, individual differences, and obstruent voicing.[10]

Figure 7 gives an overview of the speech rate (expressed in syllables per second) realized by the 10 participants in the 200 utterances. As appears from the boxplots, there are considerable differences between the single speakers, ranging from the slowest mean in Raf (3.98 syllables per second) to the highest mean in Mir (4.86 syllables per second). In addition, the intrasubject variation of speech rate varies noticeably from one speaker to another, if we compare the standard deviation of Lis (0.49) with that of Ser (0.82). An ANOVA does indeed yield a significant effect of the factor "speaker" on speech rate: $F(9, 190) = 4.07$, $p < 0.0001$.

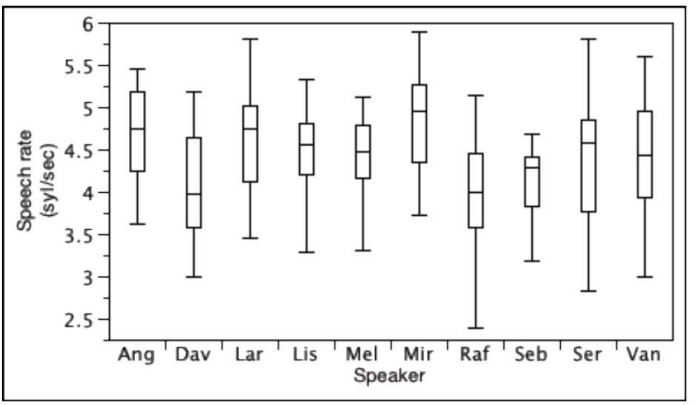

Figure 7. Speech rate (syllables per second) of the 10 speakers.

But may one also detect a correlation between speech rate and the realization of voiced obstruents? There appears to be a very slight tendency towards a higher

[10] The suggestion that speech rate should also be considered has been made by an anonymous reviewer whom I would like to thank here.

percentage of voiced obstruents when the sentences are read with a slower speech rate, as may be inferred from the coefficient of determination (R^2 = 0.109). Similarly, the percentage of unvoiced obstruents decreases with a slower speech rate, but the correlation is even weaker (R^2 = 0.017). On the whole, therefore, it is not surprising that two ANOVA did not yield significant effects for speech rate, neither on the percentage of voiced realizations nor on the percentage of unvoiced realizations.

3.6. Comparison with a native speaker

In order to compare the French pronunciation of our learners with the speech of a native speaker, the same sentences were read by a sixty-year-old Frenchman, who at the time of the recording (March 2010) worked as a language teacher at the University of Zurich. The recording took place in a sound-proof booth at the Phonetics Laboratory, employing a Neumann KM 140 microphone with a cardioid directional pattern, a frequency range of 20–20,000 Hz, and a sensitivity at 1 kHz into 1 kohm of 15 mV/Pa; the microphone was connected via a Digi002 audio interface to a computer, where the signal was processed by means of the ProTools LE 7 software (sampling rate: 44.1 kHz; quantisation: 16 bit).

Not surprisingly, the pronunciation of our native speaker (Duv) reflects the norms of standard French to a very high degree. All obstruent phonemes are realized as 100% voiced over the whole duration of the segment, at least in the contexts where this is required by the rules of French phonology, i.e. ##_V, V#_V, V_V, V#C[+voice], [_##]. The only deviations with regard to our expectations concern the context _#[−voice], where a postlexical assimilation rule predicts the devoicing of the preceding obstruent (cf. 1.2): out of six tokens pronounced by the speaker Duv, one is completely voiced, and two tokens display a partial voicing of 14% and 40% of the whole duration of the segment.

The attempt to sound as standard-like as possible is achieved, among other things, through a strategy of schwa epenthesis. A typical feature of French is the so-called *e muet* or *e caduc* (i.e., "mute" or "dropped" *e*): the schwa phoneme is often dropped in running speech, as long as the resulting consonant clusters do not exceed the number of two segments (cf. Léon 2007: 211–237). Schwa dropping is particularly common in word-final position, considering that dictionaries list the citation form of a noun like *robe* 'dress' as /ʀɔb/. Instead, our native speaker carefully realizes many schwas, also in word-final and utterance-final position.

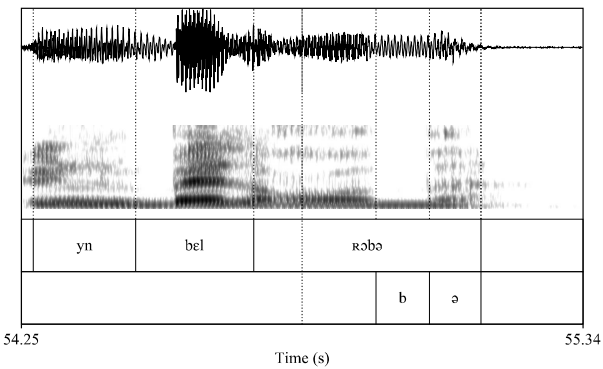

Figure 8. Word- and utterance-final schwa as pronounced by the native speaker Duv.

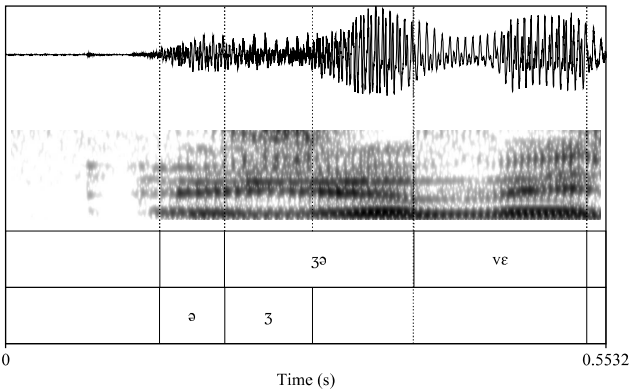

Figure 9. Utterance-initial prostethic schwa as pronounced by the native speaker Duv.

In our corpus, for instance, the noun phrase *une belle robe* 'a nice dress' occurs in utterance-final position; the speaker Duv clearly pronounces a final schwa, as becomes evident from the waveform, the spectrogram, and the TextGrid illustrated in Figure 8. In this case, the full realization of schwa constitutes not only a sociolectal feature which generally characterizes the more formal registers of French (cf. Léon 2007: 216–217), but it also favors the voiced pronunciation of /b/ which is prevented from undergoing devoicing by the presence of the paragogic vowel.

Quite interestingly, our speaker employs an articulatory strategy of vowel insertion not only utterance-finally, but also in utterance-initial position. Without being aware of the goal of the recording, Duv starts reading the first sentence of the reading list and inserts a prostethic schwa of 60 ms before the verbal phrase *je vais* 'I go' (Figure 9). Note that, quite differently from the utterance-final schwa in the noun phrase *belle robe* (Figure 8), this sort of "excrescent vowel" (Bagemihl 1991: 600) is a low-level production strategy, which obviously favors the voicing of /ʒ/; but it must be stressed that schwa prosthesis does not form part of the "Phonology" of French.

4. Concluding remarks

Obviously, the relatively limited amount of data analyzed in this preliminary study only permits us to draw a few tentative conclusions. Nevertheless, a few comments on the results presented in the subsections 3.1–3.5 are in order, particularly with regard to the claims stated in 1.4. These hypotheses were first derived from a contrastive analysis between the two obstruent systems of French and Swiss German dialects, but they also gain some plausibility from models of non-native speech perception and second language speech learning. Despite the prevalently observational character of this study, a brief discussion of these models – i.e., the "Speech Learning Model" (SLM), the "Perceptual Assimilation Model" (PAM), the "Ontogeny Philogeny Model" (OPM), and the "Markedness Differential Hypothesis" (MDH) – has been conducted in subsection 1.3 (in particular with regard to the empirical issue at stake in this study and in relation to the hypotheses to be further developed in 1.4), so there is no use in repeating it here.

Instead, we would like to maintain that the results emerging from our exploratory study already provide a piece of empirical evidence for the two main hypotheses. As regards hypothesis (i), the production data presented in 3.1–3.3 reveal a certain tendency among Swiss German learners to assimilate the [±voice] contrast of French obstruents to the native [±tense] contrast (quite differently from the native speaker analyzed in 3.6, who produced a 100% of voiced obstruents in the contexts where this is required from the pronunciation norm of standard French). Concerning hypothesis (ii), this "equivalence classification" is determined to some extent by the phonotactic contexts in which the segments occur, considering that devoicing is more frequent before another obstruent (as a result of L1 interference) and in prepausal contexts (a universal preference of sound structure).

Based on the cross-linguistic frequency of certain segment types, two additional claims regarding manner and place of articulation of voiced obstruents were formulated in 1.4. Only hypothesis (iv) is now borne out by the results presented in 3.3: it is true that /g/ undergoes devoicing more often than the other two voiced plosives. Conversely, hypothesis (iii) is not supported by our data: one cannot say that fricatives are less voiced than stops – on the contrary, /v/ is the obstruent which was most frequently pronounced as voiced by our speakers. As regards the fifth hypothesis stated in 1.4, we noted a positive transfer effect on French pronunciation deriving from the early acquisition of a language that has voiced obstruents: the three Swiss German/Italian bilinguals achieved a relatively high score of obstruent voicing compared to the other participants in the study. Along these lines, a negative transfer effect could also be detected in the Swiss German/Finnish bilingual, who had the lowest voicing score among the ten speakers (remember that Finnish almost lacks phonemic voiced obstruents).

Needless to say, many questions still remain. For instance, the examination of speech rate (cf. Section 3.5) did not yield conclusive results, even if a slight correlation between slow speech rate and a higher amount of voiced tokens could be detected (which nevertheless did not turn out to be statistically significant). Future studies could test the effect of speech rate on voicing separately for different phonotactic contexts: it might be the case that a faster speech rate enhances tendencies already emerging in the different contexts (thus voiced obstruents being more frequent in the intervocalic position and devoiced tokens occurring more often in the prepausal context). Further directions of research would include an expansion of the empirical basis, including more speakers and more tokens of the different segments; a comparison between read and spontaneous speech also constitutes a desideratum. Furthermore, it would be interesting to study the impact of obstruent voicing on the perception of both native and non-native listeners (i.e. the degree to which an unvoiced or partially voiced realizations of voiced phonemes contribute to the perception of a "foreign accent").

One last remark concerns the possible pedagogical fall-out of this kind of research. In our case, the participants have been given the opportunity to reflect, from a metalinguistic point of view, on the difficulty of pronouncing voiced obstruents for a speaker of Swiss German. A couple of weeks after the recordings, they attended a lesson which presented basic notions of articulatory and acoustic phonetics, with a particular focus on vocal fold vibration and periodic sound waves. During the lesson, they were also confronted with some audio examples from their own pronunciation and received practical hints on how to produce a

voiced obstruent (including the strategy of schwa insertion). Obviously, in order to obtain a sustained improvement of the students' L2 pronunciation, a syllabus of practical exercises would have to be built, including repeated measures of pronunciation assessment. But this raises new issues and new questions which would better be tackled in another study.

REFERENCES

Bagemihl, B. 1991. "Syllable structure in Bella Coola". *Linguistic Inquiry* 22. 589–646.

Best, C. 1995. "A direct realist view of cross-language speech perception". In: Strange, W. (ed.), *Speech perception and linguistic experience. Issues in cross-language research*. Timonium, MD: York Press. 171–204.

Best, C. and M. Tyler 2007. "Nonnative and second-language speech perception: Commonalities and complementarities". In: Bohn, O. and M. Munro (eds.), *Language experience in second language speech learning: In honor of James Emil Flege*. Amsterdam: Benjamins. 13–33.

Birdsong, D. 2007. "Nativelike pronunciation among late learners of French as a second language". In: Bohn, O. and M. Munro (eds.), *Language experience in second language speech learning: In honor of James Emil Flege*. Amsterdam: Benjamins. 100-116.

Boersma, P. and D. Weenink. 2010. Praat: doing phonetics by computer. < http://www.fon.hum.uva.nl/praat/> (Last accessed 23 Mar 2012.)

Bongaerts, T. 1999. "Ultimate attainment in L2 pronunciation: The case of very advanced late L2 learners". In: Birdsong, D. (ed.), *Second language acquisition and the Critical Period Hypothesis*. Mahwah, NJ: Erlbaum. 133–159.

CEFR 2001. *Common European Framework of Reference for Languages: Learning, Teaching, Assessment*. Strasbourg: Council of Europe.

D'Apolito, S. and B. Gili Fivela. 2009. "Il francese come lingue straniera: l'assimilazione del tratto di sonorità". In: Romito, L., V. Galatà and R. Lio (eds.), *La fonetica sperimentale. Metodo ed applicazioni*. Torriana: EDK. 216–241.

Darcy, I. and F. Kügler. 2007. "Phonological context effects for voicing and devoicing in French". *Proceedings of the 16th ICPhS*, Saarbrücken. 1257–1260.

Donegan, P. and D. Stampe. 1979. "The study of Natural Phonology". In: Dinnsen, D. (ed.), *Current approaches to phonological theory*. Bloomington: Indiana University Press. 127–173.

Eckman, F. 1977. "Markedness and the contrastive analysis hypothesis". *Language Learning* 27. 315–330.

Eckman, F. 2008. "Typological markedness and second language phonology". In: Hansen Edwards, J. and M. Zampini (eds.), *Phonology and Second Language Acquisition*. Amsterdam: Benjamins. 95–115.

Flege, J. 1987. "The production of 'new' and 'similar' phones in a foreign language: Evidence for the effect of equivalence classification". *Journal of Phonetics* 15. 47–65.

Flege, J. 1995. "Second language speech learning. Theory, findings, and problems". In: Strange, W. (ed.), *Speech perception and linguistic experience: Issues in cross-language research*. Timonium, MD: York Press. 233–277.
Fleischer, J. and S. Schmid. 2006. "Zurich German". *Journal of the International Phonetic Association* 26. 243–253.
Fougeron, C. and S. Smith. 1999. "French". *Handbook of the International Phonetic Association*. Cambridge: Cambridge University Press. 78–81.
Gass, S. and L. Selinker. 2008. *Second language acquisition. An introductory course*. (3rd ed.) London: Routledge.
Hilty, G. and J. Wüest. 1985. *Langue française: Phonétique, morphologie, syntaxe, différences de structure entre le français et l'allemand*. Zürich: Romanisches Seminar.
Horner, A. 1989. Étude de problèmes phonétiques des Zurichois apprenant le français. (Mémoire de licence, Université de Strasbourg.)
Hyltenstam, K. 1990. "Typological markedness as a research tool in the study of second language acquisition". In: Dechert, H. (ed.), *Current trends in European second language acquisition research*. Clevedon: Multilingual Matters. 23–36.
Jakobson, R. and M. Halle 1964. "Tenseness and laxness". In: Abercrombie, D. et al. (eds.), *To honour Daniel Jones*. London: Longman. 96–101.
James, C. 1980. *Contrastive analysis*. London: Longman.
Kaglik, A. and P. Boula de Mareüil. 2010. "Polish-accented French in perception and production: Transfer or universal acquisition process?" *Proceedings of Speech Prosody 2010*. Chicago. P1.24
Kohler, K. 1995. *Einführung in die Phonetik des Deutschen*. (2nd ed.) Berlin: Schmidt.
Kraehenmann, A. 2003. *Quantity and prosodic asymmetries in Alemannic. Synchronic and diachronic perspectives*. Berlin: de Gruyter.
Künzel, H. 1977. *Signalphonetische Untersuchungen deutsch-französischer Interferenzen im Bereich der Okklusive*. Frankfurt am Main: Peter Lang.
Lado, R. 1957. *Linguistics across cultures. Applied linguistics for language teachers*. Ann Arbor: University of Michigan Press.
Larsen-Freeman, D. and M. Long. 1991. *An introduction to second language acquisition research*. London: Longman.
Leemann, A. 2011. "Einfluss der Schweizerdeutschen Phonologie auf die Stimmhaftigkeit von Frikativen im L2-Englischen". Poster presented at P&P7, Osnabrück, 7 Oct 2011.
Léon, P. 2007. *Phonétisme et prononciation du français*. (5th ed.) Paris: Armand Colin.
Liberman, A., K. Harris, H. Hoffman and B. Griffith. 1957. "The discrimination of speech sounds within and accross phoneme boundaries". *Journal of Experimental Psychology* 54. 358–368.
Maddieson, I. 1984. *Patterns of sounds*. Cambridge: Cambridge University Press.
Maddieson, I. 2011. "Phonological complexity in linguistic patterning". *Proceedings of the 17th International Congress of Phonetic Sciences*, Hong Kong. 28–34.
Major, R. 2001. *Foreign accent. The ontogeny and philogeny of second language acquisition*. Mahwah, NJ: Lawrence Erlbaum Associates.
Manno, G. 2003. "The French language inside and outside Suisse Romande: Between tradition and innovation". In: Charnley, J. and M. Pender (eds.), *Living with languages. The contemporary Swiss model*. Bern: Peter Lang. 33–51.

Manno, G. 2007. "La situation du français en Suisse: Considérations démolinguistiques et de politique linguistique". *Publifarum: Constellations francophones* 7. <http://publifarum.farum.it/ezine_articles.php?art_id=45>

Nocchi, N. and S. Schmid. 2006. "Labiodentale Konsonanten im Schweizerdeutschen". In: Klausmann, H. (ed.), *Raumstrukturen im Alemannischen*. Bregenz: Neugebauer. 25–35.

Ohala, J. 1992. "What's cognitive, what's not, in sound change". In: Kellermann, G. and M. Morrissey (eds.), *Diachrony within synchrony: Language history and cognition*. Frankfurt: Peter Lang. 309–355.

Rigault, A. 1970. "L'assimilation consonantique de sonorité en français: Étude acoustique et perceptuelle". *Proceedings of the 6th International Congress of Phonetic Sciences*, Prague. 763–766.

Rutherford, W. (ed.) 1984. *Language universals and second language acquisition*. Amsterdam: Benjamins.

Schmid, S. 1997. "The Naturalness Differential Hypothesis: Cross-linguistic influence and universal preferences in interlanguage phonology and morphology". *Folia linguistica* 31. 331–348.

Schmid, S. 2005. "Spelling and pronunciation in migrant children: The case of Italian-Swiss German bilinguals". In: Cook, V. and B. Bassetti (eds.), *Second language writing systems*. Clevedon: Multilingual Matters. 184–211.

Schmid, S. 2010. "Voiced obstruents in L2 French: the case of Swiss German learners". In: Dziubalska-Kołaczyk, K., M. Wrembel and M. Kul (eds.), *New Sounds 2010. Proceedings of the 6th International Symposium on the Acquisition of Second Language Speech*. 403–408.
< http://ifa.amu.edu.pl/newsounds/> (Last accessed 25 Nov 2012.)

Schubiger, M. 1983. "Alemanic English". In: Allan, J. and B. Kettemann (eds.), *Dialektphonologie und Fremdsprachenerwerb* [Dialect phonology and foreign language acquisition]. Tübingen: Narr. 24–31.

Seliger, H. and E. Shohamy. 1989. *Second language research methods*. Oxford: Oxford University Press.

Sievers, E. 1876. *Grundzüge der Lautphysiologie zur Einführung in das Studium der Lautlehre der indogermanischen Sprachen*. Leipzig: Breitkopf & Härtel.

Skehan, P. 1989. *Individual differences in second-language learning*. London: Arnold.

Snoeren, N. and J. Segui. 2003. "A voice for the voiceless: Voice assimilation in French". *Proceedings of the 15th International Congress of Phonetic Sciences*. Barcelona. 2325–2328.

Suomi, K., J. Toivanen and R. Ylitalo. 2008. *Finnish sound structure. Phonetics, phonology, phonotactics and prosody*. Oulu: University of Oulu.

Tarone, E. 1988. *Variation in interlanguage*. London: Arnold.

Vieru-Dimulescu, B. and P. Boula de Mareüil. 2006. "Perceptual identification and phonetic analysis of 6 foreign accents in French". *Proceedings of Interspeech 2006*, Pittsburgh. 441–445.

Vieru-Dimulescu, B., P. Boula de Mareüil and M. Adda-Decker. 2007. "Characterizing non-native French accents using automatic alignment". *Proceedings of the 16th International Congress of Phonetic Sciences*, Saarbrücken. 2217–2220.

Willi, U. 1996. *Die segmentale Dauer als phonetischer Parameter von 'fortis' und 'lenis' bei Plosiven im Zürichdeutschen: Eine akustische und perzeptorische Untersuchung*. Stuttgart: Steiner.
Winteler, J. 1876. *Die Kerenzer Mundart des Kantons Glarus in ihren Grundzügen dargestellt*. Leipzig: Winter.

APPENDIX: THE 20 FRENCH SENTENCES READ BY THE PARTICIPANTS

Orthographic representation	Phonemic transcription	English translation
1. Je vais souvent à la plage.	ʒə vɛ suvɑ̃ a la plaʒ	'I often go to the beach.'
2. Maintenant je dois manger quelque chose.	mɛ̃tnɑ̃ ʒə dwa mɑ̃ʒe kɛlkə ʃoz	'Now I need to eat something.'
3. Où se trouve l'auberge de la jeunesse?	u sə tʀuv l obɛʀʒə də la ʒœnɛs	'Where is the Youth Hostel?'
4. Voulez-vous boire un café?	vule vu bwaʀ œ̃ kafe	'Do you want to drink a coffee?'
5. Beaucoup de Suisses sont partis en vacances.	boku də sɥis sɔ̃ paʀti ɑ̃ vakɑ̃s	'Many Swiss have left for a holiday.'
6. Dans la maison, il fait froid.	dɑ̃ la mɛzɔ̃ il fɛ fʀwa	'In the house, it is cold.'
7. Pierre mange très vite.	pjɛʀ mɑ̃ʃ tʀɛ vit	'Peter eats very quickly.'
8. J'adore les vagues de la mer.	ʒ adɔʀ le vag də la mɛʀ	'I adore the waves of the sea.'
9. Le garçon aide son père.	lə gaʀsɔ̃ ɛt sɔ̃ pɛʀ	'The boy helps his father.'
10. Céline aide les amis.	selin ɛd lez ami	'Céline helps the friends.'
11. Il faut qu'on l'aide.	il fo k ɔ̃ l ɛd	'He needs to be helped.'
12. C'est une blague.	s ɛt ynə blag	'This is a joke.'
13. En Normandie il y a des vagues très hautes.	ɑ̃ nɔʀmɑ̃di il j a de vag tʀɛ ot	'In Normandy, there are very high waves.'
14. Véronique porte une belle robe.	veʀɔnik pɔʀt yn bɛl ʀɔb	'Véronique wears a nice dress.'
15. François a acheté une chemise très chère.	fʀɑ̃swa a aʃte unə ʃmis tʀɛ ʃɛʀ	'Frank has bought a very expensive shirt.'

16. On ne peut pas partir à cause de la grève.	õ nə pø pa paʀtiʀ a koz də la gʀɛv	'One cannot leave because of the strike.'
17. Gardez vos bagages!	gaʀde vo bagaʒ	'Keep your bagage!'
18. Le dimanche on se lève très tôt.	lə dimɑ̃ʃ õ sə lɛf tʀɛ to	'On sunday, one gets up very early.'
19. Amélie a acheté une robe très chère.	ameli a aʃte yn ʀɔb tʀɛ ʃɛʀ	'Amélie has bought a very expensive dress.'
20. Ce soir elle porte une robe verte.	sə swaʀ ɛl pɔʀt yn ʀɔb vɛʀt	'Tonight she wears a green dress.'

Address correspondence to:
Stephan Schmid
Phonetisches Laboratorium der Universität Zürich
Rämistrasse 71
CH-8006 Zürich
Switzerland
schmidst@pholab.uzh.ch